



**Universiteit
Leiden**
The Netherlands

A compass towards equity: a data analysis framework to capture children's behaviour in the playground context

Nasri, M.

Citation

Nasri, M. (2024, December 3). *A compass towards equity: a data analysis framework to capture children's behaviour in the playground context*. Retrieved from <https://hdl.handle.net/1887/4170540>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4170540>

Note: To cite this publication please use the final published version (if applicable).

CHAPTER 3

A Novel Metric to Measure Spatio-temporal Proximity: A Case Study Analyzing Children's Social Network in Schoolyards

The contents of this chapter are based on the following publication:

- M. Nasri, M. Baratchi, Y.-T. Tsou, S. Giest, A. Koutamanis, and C. Rieffe, "A novel metric to measure spatio-temporal proximity: a case study analyzing children's social network in schoolyards," *Applied Network Science*, vol. 8, no. 1, p. 50, 2023 DOI: [10.1007/s41109-023-00571-6](https://doi.org/10.1007/s41109-023-00571-6).

Abstract

The present study aims to infer individuals' social networks from their spatio-temporal behavior acquired via wearable sensors. Previously proposed static network metrics (e.g., centrality measures) cannot capture the complex temporal patterns in dynamic settings (e.g., children's play in a schoolyard). Moreover, existing temporal metrics overlook the spatial context of interactions. This study aims first to introduce a novel metric on social networks in which both temporal and spatial aspects of the network are considered to unravel the spatio-temporal dynamics of human behavior. This metric can be used to understand how individuals utilize space to access their network, and how individuals are accessible by their network. We evaluate the proposed method on real data to show how the proposed metric impacts the performance of a clustering task. Second, this metric is used to interpret interactions in a real-world dataset collected from children playing in a playground. Moreover, by considering spatial features, this metric provides unique knowledge of the spatio-temporal accessibility of individuals in a community, and more clearly captures pairwise accessibility compared with existing temporal metrics. Thus, it can facilitate domain scientists interested in understanding social behavior in the spatio-temporal context. Furthermore, We make our collected dataset publicly available for further research.

3.1 Introduction

Increasing public awareness about the potential impact of social networks on individuals and society has led to an interest in understanding how these networks function, and how individuals are positioned in their social contexts. In an effort to investigate the structure of these networks, a considerable body of research literature has developed around the theme of social network analysis (SNA) [92–95]. The availability of wearable technologies that allow for the collection of fine-grained data from daily activities, advances in data analysis techniques and computational tools, and the need to understand the dynamics of social networks in various fields such as sociology, sports science, and healthcare, have all contributed to increased interest in SNA. SNA involves a collection of tools that adopt graph theory to perform extensive computational techniques that ultimately identify connection patterns among members of a community [96]. For instance, a static representation graph accumulates relationships between individuals in an episode, calculates static metrics such as centrality measures (e.g., closeness centrality [97]), and estimates the role of any given node (or individual) in the static representation of that network.

However, a static representation graph eliminates all temporal details, such as the frequency of events and the time difference between subsequent events, when this temporal information might include important information about the nature of the interactions at stake [98]. It also does not take into account any spatial relationships featured within those interactions.

Regarding omitted temporal detail, Figure 3.1 demonstrates the limitations of existing approaches. Throughout this paper, we will refer back to this running example - inspired by an in vivo social scenario - to illustrate our proposed method. The example data here reflect the play of four children (A, B, C, and D) during recess in a playground that features four distinct play areas: a sandpit, a bench, a combination of green areas, and a walking path, plus the entrance to the playground. In this scenario, Child-A played in the sandpit during the whole break, whilst Child-B first played in the sandpit, then engaged in a conversation with a peer on a bench, and later walked to the entrance of the playground. According to the data source in Figure 3.1-a, Child-A interacted with peers only at the beginning of recess time (i.e., T1, T2, T3) in the sandpit, where most of the peers appear to have been located during the same period, whereas Child-B interacted with peers across the entire recess time (i.e., T1, T4, T6) in different locations (i.e., sandpit, bench, and the entrance). These facts reflect distinctions that are relevant to social network structure. Yet although Child-A and Child-B had different behavioral patterns during recess, their positions in the static graph representation, as shown in Figure 3.1-b, appear the same: the graph shows an equal number of edges (partners) with equal weights (number of interactions), which yields the same representation as in their static metrics, such as centrality measures. This example illustrates how these static metrics are not sensitive to temporal changes, and fail to capture complex dynamics that could be contained within such data.

To tackle this problem, Kostakos [98] proposed a “temporal” graph representation, in which the relationship between nodes and the position of each node was investigated in the temporal context of an entire network, rather than in an aggregated format. Based on this representation, Kostakos [98] introduced new metrics, including average geodesic proximity and average temporal proximity, which consider edge availability over time and take into account possible wait times for each node before meeting the next node. Specifically, these are measures of how quickly an individual is accessible by their network, and how quickly the network is accessible by the individual. Figure 3.1-c does show this temporal representation: in contrast to a static analysis, Child-A and Child-B are positioned differently in this representation. Here, the temporal metrics preserve information about the order and the frequency of events over time.

Yet, even in this temporal representation of a network, one aspect that is still

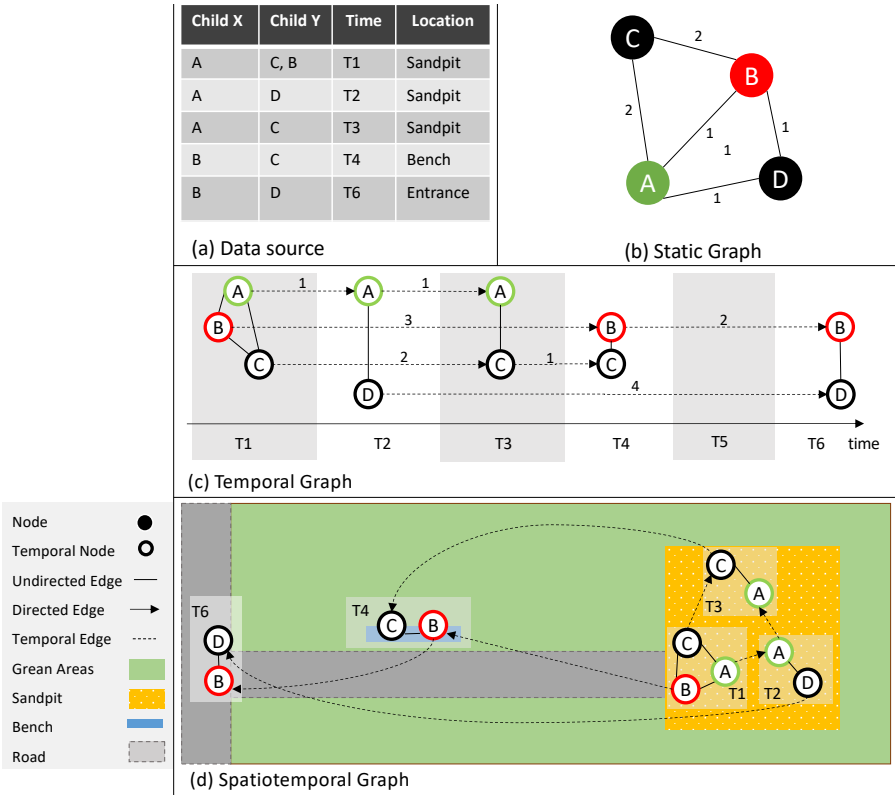


Figure 3.1: Visual comparison on how a data source of interactions between four children in a playground is represented via b static graph, c temporal graph, and d Spatio-temporal graph. While Child-A and Child-B have different temporal patterns (e.g., frequency of interactions and difference in time between interactions) and spatial features (e.g., location of interactions) in the (a) data source, their position in the (b) static graph is the same. The (c) temporal graph could better preserve temporal details, but the spatial information is ignored. And finally, the (d) spatio-temporal representation could preserve the temporal and spatial information of interactions.

entirely overlooked is that of the physical environment. In the context of the example illustrated in Figure 3.1, investigating the temporal accessibility of a child in the playground does not inform us about how the child utilized the environment: did the child walk around and interact with peers (e.g., Child-B in Figure 3.1-d), or simply

remain in a crowded spot while making contact with peers (e.g., Child-A playing in the sandpit in Figure 3.1-d)? To understand the impact of the physical environment on an individual's interactions with their network, it is crucial to consider spatial features in analyzing the social network. The temporal metric estimates availability only by considering temporal differences between nodes. The spatial aspect of this availability, i.e., the location of contacts in the physical environment, is still ignored in a temporal graph representation.

The literature has already documented the importance of considering spatial context when examining social networks. Several studies have examined the effect of environments on human behavior, for example, the impact of physical and social environments of neighborhoods on child and family well-being [99], the impact of office design on employee's office usage and social interactions [100], and the impact of schoolyard design on children's movement and their social behavior [1]. In fact, including the spatial element of social networks enables us to understand better how the physical environment impacts individuals' behavior [101] and how users utilize the environment to interact with their network.

While the impact of the physical environment on social networks has been confirmed in various studies, limited attention has been given to the integration of these dimensions into a unified framework. Consequently, there is a critical need for research that addresses this gap by introducing spatio-temporal metrics capable of quantifying the complex interplay between social and physical environments.

In order to examine the interdependency between spatial and social contexts of interactions and continue to build a better understanding of the impact of the physical environment on social networks, the present study pursues two main aims. First, we introduce a new metric for measuring the spatial and temporal dynamics of social networks, known as average spatio-temporal proximity. By considering the spatial differences between temporal networks, our proposed metric examines how spatially sparse these networks are and how spatio-temporally a user is connected to its network. Specifically, this metric calculates the shortest path based on the spatial distance of temporal events and estimates how far, on average, nodes (or individuals) in a network need to move to reach a specific node, and how far, on average, a node needs to move to reach the rest of the network. Such a metric would enable a more comprehensive understanding of the factors influencing social accessibility (i.e., examining how the physical and social context reciprocates around individuals).

Second, we evaluate the proposed metric in the context of a case study whose goal was to investigate social interactions among a group of 32 children playing in a schoolyard during recess. By measuring the spatio-temporal proximity of individuals over time, we aim to better understand the social behavior of these children and

their interactions within the context of their spatial environment.

Overall, this paper makes the following contributions:

- We introduce a new metric to quantify the proximity of nodes in a graph that considers both temporal and spatial characteristics of contacts in order to measure how individuals in a network utilize space to make contacts.
- We present and make publicly available a new dataset collected from children's interactions and movements in a schoolyard during recess. This dataset uniquely registered face-to-face contacts and children's locations over time, measured via wearable proximity sensors and GPS loggers, respectively. This dataset is available in a public repository for future use in spatio-temporal research.
- We evaluate the performance of our proposed metric to show how it impacts the performance of a clustering task in identifying clusters in a network, compared to temporal measures alone.
- We further demonstrate how this metric can be used to illuminate the underlying interactions in our Schoolyard Dataset.

The rest of this paper is organized as follows: In [Related work](#), we discuss the literature relevant to our study. The [Problem Definition](#) introduces the computational problem at stake in our proposed methodology. [Methodology](#) presents the details of our method. The experimental settings and results are presented in [Experiments](#). Finally, [Conclusions](#) summarizes the study and discusses future research directions.

3.2 Related work

To provide relevant context, here we reference further literature on two main approaches to the fusion of spatial information with time-series data: location-based social network (LBSN) analysis, where spatial information is intertwined with social networks, and spatio-temporal pattern mining, where relevant research is scanned to identify spatio-temporal features, as a broad concept.

3.2.1 Location-Based Social Network (LBSN) Analysis

Most studies in LBSN analysis arose from the urban computing field and social media networks analysis mainly due to the availability of large-scale databases such as Foursquare, Twitter, Instagram, and Google Places [102–105]. In this body of

research, the interdependency between spatial and social contacts lies in the co-presence of two persons in the same physical locations, or the sharing of a similar location history, common behavior, or activities. In these cases, interdependency is implied from users' location data without assessing face-to-face interactions from individuals' perspectives [106]. The aim of this line of research is often to design physical location (or activity) recommendation systems [107–109], travel planning [110, 111], and marketing [112] from accumulated spatial behavior, mainly by considering similarity among users. In contrast, spatial information in the present study is used as a descriptive machine learning model for the purpose of improving the performance of a downstream task and for discovering more about the nature of interactions rather than for predicting social contacts or common interests.

3.2.2 Spatio-temporal Pattern Mining

Recently, interest has been growing in areas related to the adoption of deep neural network models to generate spatio-temporal features in varied applications. These include crowd flow prediction [113], social event detection [3, 114], and financially aware social network analysis [115]. Although the performance of deep models for extracting spatio-temporal features has appeared promising, these models require a huge amount of representative data to learn about generalizable spatio-temporal embeddings. Therefore, these methods are hardly applicable for small datasets in which the duration of data assessments and the number of users is limited. In addition, collecting large-scale data from certain populations (e.g., patients in a hospital or children in a playground) is often impossible because of privacy protection concerns. Therefore, analyzing spatial patterns over time via classical pattern mining techniques has been the focus of many studies in domains such as animal revisitation analysis [116], disease spread patterns [117], and the tourism industry [118, 119]. This line of research often involves extracting time domain features (e.g., residence time, the time between visits) and applying a downstream task to the extracted features, such as clustering and prediction.

Analyzing temporal events via graphs, as in a “temporal graph”, was first introduced by Kostakos [98]. Using this representation, temporal metrics were defined to assess geodesic and temporal proximity. These metrics estimated how quickly a node was accessible by the network, and how quickly a network was accessible by a node. Accessibility in temporal proximity was defined as the time difference between temporal events in a weighted shortest-path algorithm. Geodesic proximity was calculated using the number of hops between temporal events in an unweighted shortest-path algorithm. Neither metric took into account the spatial context of temporal events, e.g., the geo-distance between nodes.

The present study is intended to further build upon the work of Kostakos [98]. Specifically, temporal metrics proposed by Kostakos [98] are further developed to incorporate the spatial aspect of social contacts and to examine the spatio-temporal accessibility of individuals within a network.

3

3.3 Problem Definition

Assume a spatio-temporal proximity dataset as a set of tuples in the form of $\langle v_i, v_j, s_t \rangle$ where $\{v_i, v_j \in \mathbf{V}\}$ are two entities in the set of entities $\mathbf{V}$ that are detected in proximity to each other, where $s_t = (x, y, t)$ represents the coordinate where the contact between these v_i and v_j is registered at time t , and $t \in \{1, \dots, T\}$ represents the timestamp of the contact. Such a dataset can be collected, for example, when the two entities are equipped with a proximity sensor (to detect contacts) and a GPS sensor, to acquire the average coordinate representing the point of contact between the entities.

First, we are interested in creating a spatio-temporal graph representation $g = \langle \mathbf{V}_{s_t}, \mathbf{E} \rangle$, where $\{v_{i,s_t} \in \mathbf{V}_{s_t}\}$ are sets of nodes each represented over timestamps t , at a location denoted by s , and $\mathbf{E}$ is a set of instant and spatio-temporal edges. Instant edges are unweighted and undirected links between spatio-temporal nodes v_{i,s_t} and v_{j,s_t} indicating contact between two distinct nodes v_i and v_j , $i \neq j$ at location s in time t . Spatio-temporal edges are weighted and directed links between consecutive pairs v_{i,s_t} and $v_{i,s_{t+1}}$ of node v_i (which is weighted based on the spatio-temporal proximity assuming that the direction of edges) represents the temporal order.

Second, we are interested in finding clusters $C_i = \mathbf{V}_{C_i}$ where $\mathbf{V}_{C_i} \subseteq \mathbf{V}$, such that all nodes $\mathbf{V}_{C_i}$ within each cluster C_i are spatio-temporally similar to each other and dissimilar to the nodes in other clusters.

3.4 Methodology

In this section, we present essential mathematical preliminaries as the foundation for understanding our proposed methodology, and the problem at stake. Then, we explain our proposed spatio-temporal graph and its proximity metrics.

3.4.1 Background and Preliminaries

Graphs are mathematical structures used to model pairwise relations between objects. In this structure, objects are represented as nodes that are connected by edges [120]. In order to preserve temporal information via graph representations, Kostakos [98] introduced the idea of temporal graphs. In this representation, contrary to static representations, the temporal relationship between nodes and the position of each node in the temporal context of the entire network is preserved. According to Kostakos [98], a temporal graph is constructed in three steps:

1. One temporal node per original node per point in time is created. Thus, node v_i in the original graph is represented as $v_{i,t}$ in the temporal graph where $t \in [1, \dots, T]$ are all the data points in which node v_i is communicated; e.g., Child-B in Figure 3.1 is represented by the set of instances $v_{B,t} = [v_{B1}, v_{B4}, v_{B6}]$.
2. Consecutive pairs $v_{i,t_x}, v_{i,t_{x+1}}$ with directed edges of weight $t_{x+1} - t_x$ for each set of instances v_i are linked, representing the temporal distance between the pair. For example, in Figure 3.1, the weight between nodes v_{B1} and v_{B4} is 3 time points.
3. The unweighted, undirected edges are used to link instantaneous communications between nodes. E.g., an interaction between Child-A and Child-B in Figure 3.1 at time t_1 is instantiated as an undirected link between Child-A and Child-B.

Therefore, each node v_i in Figure 3.1-b is converted to a directed chain of nodes $v_{i,t}$ that represent all temporal instances of nodes over time. Following this protocol, the temporal graph is generated as shown in Figure 3.1-c.

Furthermore, Kostakos [98] defined average geodesic proximity $G(v_i, v_j)$ on a temporal graph as the measure of “on average, how many hops is node v_i away from node v_j ”. This metric is formulated as follows:

$$G(v_i, v_j) = \frac{1}{n} \sum_{t=1}^T g(v_i, v_j, t), \quad (3.1)$$

where $g(v_i, v_j, t)$ is the shortest path, starting from node v_i at time t to the most accessible version of node v_j over time. $t \in [1, \dots, T]$ is the set of time points that node v_i is active (or communicating). n is the number of finite $g(v_i, v_j, t)$ values, e.g., $g(A, D, t=1) = 2$ (i.e., v_{A1}, v_{A2}, v_{D2}) in our example in Figure 3.1. In this metric, the lowest number of hops defines the most accessible instance.

Consequently, the $G_{in}(v_i)$ and $G_{out}(v_i)$ are calculated for node v_i as formulated in Equation 3.2:

$$\begin{aligned} G_{in}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^M G(v_j, v_i), \\ G_{out}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^M G(v_i, v_j) \end{aligned} \quad (3.2)$$

Where $G_{in}(v_i)$ and $G_{out}(v_i)$ are measures of “on average, in how many hops is v_i reached by the rest of the network” and “on average, in how many hops does v_i reach the rest of the network”, respectively. n is the number of finite $G(v_j, v_i)$ and $G(v_i, v_j)$ values, and $v_j \in \mathbf{V} = \{v_1, v_2, \dots, v_M\}, v_j \neq v_i$ is sets of nodes in a network with M nodes.

Similarly, average temporal proximity $P(v_i, v_j)$ is the measure of, “on average, the time it takes to go from v_i to v_j ” and is formulated as follows:

$$P(v_i, v_j) = \frac{1}{n} \sum_{t=1}^T p(v_i, v_j, w_t, t), \quad (3.3)$$

where $p(v_i, v_j, w_t, t)$ is the weighted shortest path, starting from node v_i at time t to the most accessible version of node v_j over time. $t \in [1, \dots, T]$ is the set of time points that node v_i is active (or communicated). n is the number of finite $p(v_i, v_j, w_t, t)$ values e.g., $p(A, D, w_t, t_1) = 1$ (i.e., v_{A1}, v_{A2}, v_{D2}) in Figure 3.1. The weight w_t equals the time difference between nodes at a given path. Consequently, the $P_{in}(v_i)$ and $P_{out}(v_i)$ are calculated as formulated in Equation 3.4:

$$\begin{aligned} P_{in}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^M P(v_j, v_i), \\ P_{out}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^M P(v_i, v_j) \end{aligned} \quad (3.4)$$

The P_{in} and P_{out} are measures of “how quickly, on average, v_i is reached by the rest of the network” and “how quickly, on average, v_i reaches the rest of the network”. n is the number of finite $P(v_j, v_i)$ and $P(v_i, v_j)$ values, and $v_j \in \mathbf{V} = \{v_1, v_2, \dots, v_M\}, v_j \neq v_i$ is sets of nodes in a network with M nodes. This metric estimates nodes’ accessibility by obtaining its availability over time.

To better understand these two temporal metrics, we calculated them in the playground example presented in the Introduction (See Figure 3.1). Table 3.1 and Table 3.2 show the result of average geodesic proximity and average temporal proximity, respectively. The results demonstrate that, despite the static measures, Child-A and Child-B have different *in* and *out* measures (both geodesic and temporal). Child-A has the lowest G_{in} and P_{in} as all peers find the child shortly at the start of the break, and the highest G_{out} due to its absence in communication during the second half of the break. This might be because the child is more involved in solitary games, rather than being with peers during recess. The lowest values belong to Child-B in both G_{out} and P_{out} because Child-B accessed peers across the break and interacted with all peers.

Table 3.1: The result of average geodesic proximity on playground example

	G				G_{out}
	A	B	C	D	
A	0	2.33	1.33	2.67	2.111
B	1	0	1	2	1.333
C	1	1.33	0	3.33	1.889
D	1	1.5	3	0	1.833
G_{in}	1	1.722	1.778	2.667	

Table 3.2: The result of average temporal proximity on playground example

	P				P_{out}
	A	B	C	D	
A	0	0.33	1.67	0.39	0.295
B	0.14	0	0.16	0.14	0.148
C	0.15	0.18	0	0.32	0.219
D	0.32	0.14	0.52	0	0.325
P_{in}	0.202	0.218	0.281	0.284	

3.4.2 Spatio-temporal Proximity

Both of the metrics that are explained in the previous section overlook spatial dependency among nodes, and consequently, ignore the spatial behavior of individuals (e.g., whether a child is interacting in different areas over the course of recess, or

the child stays in a single spot and interacts with peers who come over to that same spot). Obtaining such detailed spatio-temporal behavior is not possible via the existing measures (e.g., average geodesic proximity and average temporal proximity). To address this gap, we introduce spatio-temporal graphs and define average spatio-temporal proximity as $S(v_i, v_j)$, which takes into account spatial dependencies among nodes. By considering the spatial distances, e.g., Euclidean distance, the spatio-temporal proximity examines how spatially close the user is to their network. To create a spatio-temporal graph, we followed the same steps as in the temporal graph protocol, except in step (2), where the consecutive pairs $v_{i,s_{tx}}, v_{i,s_{tx+1}}$ are linked with the directed edges of weight $d(v_{i,s_{tx}}, v_{i,s_{tx+1}})$, which is the spatial distance between instances $v_{i,s_{tx}}$ and $v_{i,s_{tx+1}}$ (instead of temporal differences).

Accordingly, average spatio-temporal proximity is calculated based on this spatio-temporal graph, and allows the weights to be the spatial distances among the edges. This metric investigates “on average, how long (spatio-temporally) it takes for the network to reach v_i (S_{in})”, and “on average, how long (spatio-temporally) it takes for v_i to reach the rest of the network (S_{out})”. The average spatio-temporal proximity is formulated as follows:

$$S(v_i, v_j) = \frac{1}{n} \sum_{t=1}^T s(v_i, v_j, w_s, t), \quad (3.5)$$

where $s(v_i, v_j, w_s, t)$ is the weighted shortest path, starting from node v_i at time t to the most accessible version of node v_j over time. $t \in [1, \dots, T]$ is the set of time points that node v_i is active (or communicated). n is the number of finite $s(v_i, v_j, w_s, t)$ values. The weight w_s is equivalent to the spatial distance between nodes v_i and v_j at a given path, assuming that the location coordinates of all the contacts are available.

Furthermore, the spatio-temporal proximity is calculated per pair of nodes in a community, and the in-degree (S_{in}) and out-degree (S_{in}) variants are calculated by taking the average over the in-reached and out-reached variables, respectively.

$$\begin{aligned} S_{in}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^M S(v_j, v_i), \\ S_{out}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^M S(v_i, v_j) \end{aligned} \quad (3.6)$$

The $S_{in}(v_i)$ and $S_{out}(v_i)$ are measures of “how long, on average, v_i is accessible by the rest of the network” and “how long, on average, it takes on average for v_i

to access the rest of the network". n is the number of finite $S(v_j, v_i)$ and $S(v_i, v_j)$ values, and $v_j \in \mathbf{V} = \{v_1, v_2, \dots, v_M\}$, $v_j \neq v_i$ is sets of nodes in a network with M nodes. Similarly, the result of this analysis for the playground example is shown in Table 3.3. To calculate this table, it is assumed that the sandpit, bench, and entrance area are located at (3, 1), (2, 2), and (1, 1) Cartesian coordinates, respectively, and that the interactions happened around these areas. As described

Table 3.3: The result of average spatio-temporal proximity on playground example

	S				S_{out}
	A	B	C	D	
A	0	0.85	0.1	0.83	0.594
B	0.141	0	0.1	0.48	0.241
C	0.1	0.43	0	1.2	0.579
D	0.1	0.7	0.2	0	0.333
S_{in}	0.114	0.66	0.133	0.84	

in Table 3.3, Child-A has the lowest S_{in} : during time points when the child was available (i.e., T1, T2, and T3), all other peers were in the same spot (sandpit) and could access the child with a short distance to travel. Meanwhile, Child-D got the highest S_{in} because the child was only available at T2 and T6, and if anyone wanted to reach Child-D after T2, they had to pass a long spatio-temporal route (i.e., temporally from T2 to T6, and spatially to the entrance, in order to find the child at the Entrance on T6). Regarding the S_{out} , Child-A has the highest score because some peers were not presented at T2 and T3 in the sandpit. Therefore, the child had to wait and reach them in different locations. In contrast, Child-B could more easily access peers (lowest S_{out}) because the child was present across recess time in different locations. Thus, our proposed metric quantifies accessibility by unifying the spatial and temporal dimension of interactions, and provides unique knowledge on how individuals and their networks utilize the spatial environment to make contacts over time.

3.5 Experiments

We investigated the performance of the proposed metric in identifying social groups, based on the most accessible peers. Specifically, we were interested in knowing whether adding spatial context to the existing temporal metrics (i.e., geodesic and temporal proximity) could help better identify groups in a network.

We evaluated the proposed methodology by utilizing a clustering performance task. Here, we first explain our experimental design and introduce the datasets adopted in our experiments. Then, the evaluation metrics are presented. Next, we show the result of our analysis in two sections: The first part evaluates the performance of a clustering task using the proposed metric, compared with the existing temporal metrics as the baseline. The second part analyses children's behavior in the schoolyard dataset.

3

3.5.1 Experiment Settings

Experiments were designed using the Python programming language. The code has been made available in a public repository¹.

The NetworkX module was used to execute social network analysis, and Dijkstra shortest-path algorithm [121] was used for the calculation of proximity metrics (i.e., geodesic, temporal, and spatio-temporal proximity) on the spatio-temporal graph. The SciPy and Scikit Learn packages were used to perform clustering tasks and evaluate their performance, respectively. Hierarchical clustering [122] was used to cluster the proximity matrices, which was a natural choice for clustering data in the form of an adjacency matrix, especially when the dataset size is small. In addition, hierarchical clustering serves best in exploratory analysis where there is no specific outcome variable defined. For this purpose, first, the Pearson correlation coefficients of the pairwise matrices were calculated separately across in-degree and out-degree components to acquire how in-degree proximity (i.e., accessing the child by the network), and out-degree proximity (i.e., accessing the network by the child), were correlated among children. Consequently, we obtained a symmetric correlation coefficient matrix (CCM) per in-degree and out-degree metrics, thereby determining the strength and direction of the pairwise relationships among children. Next, the upper triangular area of the in-degree CCM was merged with the lower triangular area of the out-degree CCM, to create a non-symmetric CCM that described both in-degree and out-degree correlations among children. Such a non-symmetric CCM was used in the hierarchical clustering algorithm to identify groups based on mostly correlational relationships. Furthermore, the obtained matrix was used in creating heatmaps to visualize the relationships between children in a more organized way, which aids in data exploration and analysis.

In designing the hierarchical clustering algorithm, three different linkage methods, including complete, average, and average group, were investigated to understand how different strategies in creating the linkage matrix impact the performance

¹<https://github.com/maaeedee/spatio-temporal-network.git>

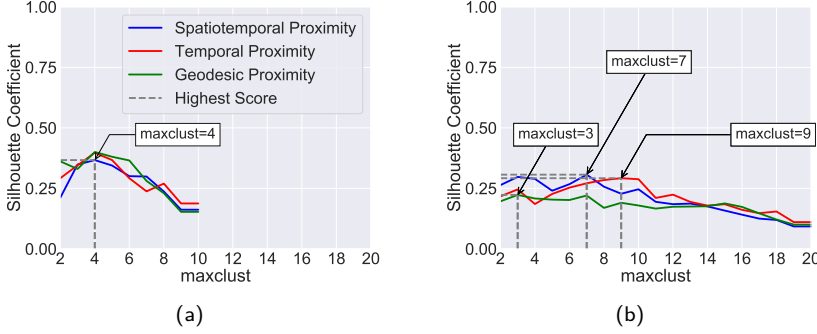


Figure 3.2: The maximum Silhouette coefficient score is used to find the optimum ‘maxclust’ value t for the hierarchical clustering algorithm on the range of $t \in [2, N]$, in which N is the total number of data points for (a) PG-1 ($N = 10$) and (b) PG-2 ($N = 20$).

Accordingly, the ‘maxclust’ is selected as 4 for all three measures in PG-1 and as $G = 9, P = 3, S = 7$ for PG-2 based on the maximum Silhouette coefficient score.

of the clustering algorithm. In short, the ‘complete’ linkage method defines the distance between two groups as the distance between the most distant pair of points, while the ‘average’ defines this distance as the average pairwise distance between data points in two groups, and ‘average group’ defines it as the average pairwise distance between all data points when two groups are merged. Moreover, the correlation distance was used to ensure that the clustering algorithm considered the strength and direction of the correlations between variables. The ‘maxclust’ was selected as the criterion, which finds a minimum threshold to ensure the cophenetic distance between any two original points in the same flat cluster is no greater than the threshold, and no greater than t flat clusters are formed. Since the number of clusters was unknown, the optimum value of t was chosen based on the result of clustering performance, using the Silhouette coefficient score [123] on the range of $t \in [2, N]$, in which N is the total number of data points. This range was chosen based on the assumption that there were at least two clusters in a playgroup, and that, in the worst-case scenario, there was no child in the same cluster with peers. Since ‘complete’ linkage, on average, performed slightly better than the other two linkage methods (as shown in Table 3.5), this linkage matrix was used to obtain the optimum number of clusters. Figure 3.2 shows results for both playgroups. Based on these figures, the ‘maxclust’ was selected as 4 for all proximity measures in PG-1 and $[G = 9, P = 3, S = 7]$ for PG-2 based on the maximum Silhouette coefficient score.

3.5.2 Schoolyard Dataset

We collected spatio-temporal data via wearable sensors from 32 children, aged between 4 and 12 years old, in two playgroups, PG-1 and PG-2, during recess in the schoolyard of a primary special education school in the Netherlands. These data are now publicly available, so other researchers may replicate our results².

Playgroups PG-1 and PG-2 were children from junior and senior grades, respectively, who were assigned to designated playgrounds. As shown in Figure 3.3, PG-1 used areas I, II, and III, which included a soccer field, bench, climbing structure, and sandpit. PG-2 spent recess time in areas IV and V, where a bench, table tennis, climbing frame, and soccer field were located. Before starting each break, children were asked to wear a belt with a mounted proximity tag and a GPS logger. The proximity tags detected face-to-face contacts, and the GPS loggers recorded the geographic locations of the users. Specifications for the datasets are described in Table 3.4. In addition to sensor data, video recordings, and field observation were conducted simultaneously to gain insight into the psychological aspects of children’s interactions and social behavior. Parental informed consent was obtained before conducting the data collection. Approval for the study was obtained from the Leiden University Ethical Committee.

Table 3.4: Specification of Schoolyard Datasets for playgroups PG-1 and PG-2

Datasets	#Groups	#SessionsxDuration	#Participants	Age (M ± SD)
PG-1	2	2 × 30min, 1 × 15min	11	5.7 ± 0.96
PG-2	3	2 × 15min	22	10.43 ± 0.77

In order to evaluate the effectiveness of our proposed method and verify the results, we need ground-truth data (i.e., unbiased and true values serving as a reference for evaluating the method). The ground-truth informs us about the true pairwise relationship among children, e.g., the true values on how often a child was in meaningful interaction with peers, where precisely in the schoolyard, and for how long. This data could be presented, for example, in the form of pairwise group membership over time by defining ‘group’ as peers who are most accessible by their group mates. Due to privacy concerns regarding any child population, existing videos associated with the data collection were recorded only from focal areas, following the approved ethics application. This made it practically impossible to create ground-truth for the collected sensor data and to systematically verify the

²<https://doi.org/10.34894/ZCPXDW>

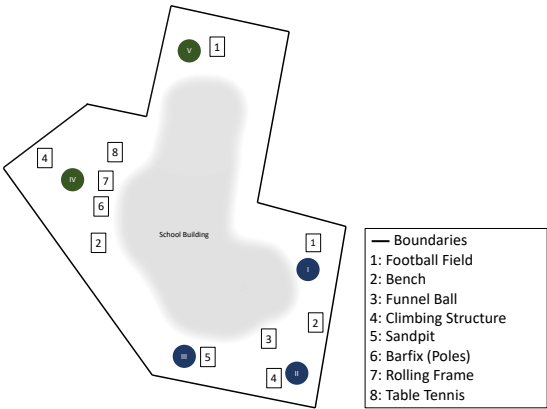


Figure 3.3: Schoolyard floor plan: Junior grades spend their recess time in areas I, II, and III, while senior grades are in areas V and IV. The schoolyard environment has different play structures, as described in the legend, which children may use during recess.

results. Moreover, evaluating the concept of pairwise proximity or accessibility over time might not be easily understood and coded by monitoring video recordings or direct observations, especially with regard to children's age groups who often tend to change their peer group and attend different types of play over the course of recess time. Yet, the collected video recordings and the field observation reports allowed us to discover certain behaviors (e.g., identifying the most visited areas, popular activities (or games) among children of a certain group, etc.), to draw a general picture of schoolyards activities, and to verify the obtained results to some extent. Therefore, the experiments in this section are at an exploratory level considering different age groups and playground designs.

3.5.2.1 Data Preparation and Pre-processing

Proximity Data: Wearable proximity tags were used to detect face-to-face contact between children during recess. If any pairs of proximity tags (i.e., children) were at a distance of up to $1.5m$ in the degree of about $60 - 70^\circ$, then they registered each other's unique code via Bluetooth. They wirelessly sent this detection to two base stations installed in two different locations in the schoolyard. The installation locations were estimated beforehand to obtain the highest coverage over the schoolyard by considering the fact that each base station covers an area of up to $25m^2$.

Hence, some children might have a low detection rate for various reasons, e.g., playing far away from the base stations or tag malfunctioning. To compensate for this error, the detection rate is calculated per child by dividing the duration of detection by beagle over the total duration of the break. Only children with a detection rate higher than 50% are included for further analysis. As a result, three children from PG-2 who had a low detection rate were excluded from further analysis. The data reported in Table 3.4 is after this exclusion. The obtained contact data is then fused with GPS data to obtain the location of interactions

GPS Locations: The GPS logger(i-gotU GT-120 USB) recorded location coordinates in Longitude and Latitude per second. The GPS locations were converted afterward to Cartesian space and merged with proximity contacts to obtain the location of interactions. As suggested in the literature [91, 124, 125], we adopted Euclidean distance to examine the spatial distances between temporal nodes (w_s in Equation 3.5). In our study, it was assumed that children in face-to-face contact are almost in the same location, to use the data more efficiently. Therefore, the available GPS data of one of the pairs in interaction was used for the other, in case the GPS data were missed. In addition, a 10-second window was used to search for the location of individuals in an interaction, and the median values were used to report the location coordinates at the given time. This merged data was used in our analysis to create the spatio-temporal graph.

3.5.3 Evaluation Metrics

This section describes the evaluation metrics used to examine the performance of a clustering task. Since no ground-truth data were available for the schoolyard dataset, we decided to use the following evaluation metrics, which do not require ground-truth but still provide a quantitative assessment of the quality of the clusters found in these datasets when different proximity metrics were used. We expect to observe clusters with higher quality when more informative proximity metrics are used.

1. **Silhouette score** [123] indicates how well clusters are separated from each other. This metric is defined as follows:

$$SC = \frac{1}{N} \sum_{i=1}^N \frac{(b_i - a_i)}{\max(a_i, b_i)}$$

where a is the mean distance between sample i and all other points in the same cluster, and b is the mean distance between sample i and all other points in the next nearest cluster. N is the total number of points.

2. **Calinski-Harabasz score** [126] is the ratio between the within-cluster dispersion and the between-cluster dispersion, and is defined as follows:

$$s = \frac{\text{tr}(B_k)}{\text{tr}(W_k)} \times \frac{n_E - k}{k - 1}$$

where n_E and k are the data and cluster sizes, respectively. $\text{tr}(B_k)$ is the trace of the between-group dispersion matrix, and $\text{tr}(W_k)$ is the trace of the within-cluster dispersion matrix defined as follows:

$$W_k = \sum_{q=1}^k \sum_{x \in C_q} (x - c_q)(x - c_q)^T, \quad B_k = \sum_{q=1}^k n_q (c_q - c_E)(c_q - c_E)^T$$

where C_q is the set of points in cluster q , with cluster center equal to c_q . c_E is the centroid of cluster E and n_q is number of points in cluster q .

3. **Davies-Bouldin score** [127] is defined as the ratio between the cluster scatter and the cluster's separation. A lower value of this metric will indicate a better clustering and it can be calculated as follows:

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} R_{ij}, \quad R_{ij} = \frac{s_i + s_j}{d_{ij}}$$

where s_i is the diameter of cluster i and d_{ij} is the distance between cluster centroids i and j . DB is the Davis-Bouldin index.

3.5.4 Results and Discussion

In this section, we report the results of our experiments in two main areas: evaluation of the performance of the proposed metric by obtaining the quality of clusters in a hierarchical clustering task (Clustering Performance Evaluation), and the in-depth analysis of our Schoolyard Datasets, to see how children interacted with peers and utilized the physical environment.

3.5.4.1 Clustering Performance Evaluation

Table 3.5 shows the results of the clustering performance task using Silhouette, Calinski-Harabasz, and Davies-Bouldin scores in clustering CCM of pairwise relation. In pairwise relations of PG-1, the geodesic proximity and temporal proximity obtained slightly higher Silhouette and Calinski-Harabasz scores, and lower

Davies–Bouldin scores than spatio-temporal proximity in the three linkage matrices. The superiority of geodesic and temporal proximity might be due to the fact that most children in PG-1 played with bikes across the playground, according to the field observations during the data collection. This activity uniformly increased face-to-face contact with all other children across the playground, thereby making them uniformly accessible. Therefore, pairwise spatial proximity might not be influential in describing pairwise proximity for PG-1. Meanwhile, in the senior group, PG-2, our proposed spatio-temporal measure and temporal measure identified clusters with higher quality than geodesic proximity by obtaining higher Silhouette and Calinski-Harabasz scores. Whilst the result was not consistent for Davies–Bouldin scores. This competitive result in temporal and spatio-temporal metrics indicates that in addition to focusing on the temporal context, including the spatio-temporal contexts of interactions contributed to describing children’s social behavior and their position in the social network in PG-2. The availability of a larger playground area, the tendency to play in group settings in certain areas in senior groups, and more interest in stationary activities such as sitting and chatting with peers might be the reasons why the spatial context provided a positive impact on the performance of the clustering algorithm of this playground, compared with PG-1.

As Table 3.5 shows, the performance of linkage methods was inconsistent across different metrics and evaluation scores. It is important to note that the performance of different linkage methods can vary depending on the specific dataset and the nature of the underlying data. Therefore, to optimize clustering performance, the linkage method could be defined as a hyperparameter and optimized specifically per dataset. In the present study, we adopted the result of the ‘complete’ linkage matrix to conduct the analysis on the schoolyard dataset in the next sections of the analysis.

Overall, our results highlight the value of including the spatial context in examining individuals’ social behavior, and in assessing their position in the network.

3.5.4.2 In-Depth analysis of Schoolyard Data

In this analysis, accessibility is defined in three dimensions: (1) *geodesic proximity* - $G(v_i, v_j)$, (2) *temporal proximity* - $P(v_i, v_j)$, and (3) *spatio-temporal proximity* - $S(v_i, v_j)$. Each dimension provides unique knowledge on how individuals and their networks were accessible in the geodesic space, over time, and in the spatio-temporal context. Therefore, “accessibility” will heretofore refer to all three of these dimensions, throughout the rest of this paper.

As previously discussed, the CCM was used to create heatmaps as depicted in Figure 3.4 for both playgroups. The x and y axis show children from different groups

Table 3.5: The result of clustering performance in pairwise relations, based on CCM of geodesic proximity $G(v_i, v_j)$, temporal proximity $P(v_i, v_j)$, and spatio-temporal proximity $S(v_i, v_j)$ for three different linkage methods: complete, average and average group. The best performance per linkage method is indicated in boldface, and the best performance among all linkage methods is shown with (*). a-group stands for Average group.

Data	Metric	Silhouette			Calinski-Harabasz			Davies-Bouldin		
		complete	average	a-group	complete	average	a-group	complete	average	a-group
PG-1	$G(v_i, v_j)$	0.400*	0.380	0.312	7.865	9.763*	6.397	0.722	0.622*	0.994
	$P(v_i, v_j)$	0.397	0.397	0.397	9.166	9.166	9.166	0.734	0.734	0.734
	$S(v_i, v_j)$	0.367	0.367	0.313	7.363	7.363	6.369	0.814	0.814	1.170
PG-2	$G(v_i, v_j)$	0.223	0.241	0.212	5.879	5.507	5.324	1.454	0.893	0.977
	$P(v_i, v_j)$	0.292	0.292	0.249	8.276*	8.276*	8.006	0.666	0.666	1.465
	$S(v_i, v_j)$	0.307	0.308*	0.263	7.438	6.713	6.637	0.608*	0.733	1.583

in this figure. Specifically, there are two groups (A and B) in PG-1 and three groups (A, B and C) that have similar break schedules, separated by green lines. The first digit of each child's identifier indicates their group. The results are shown as a heatmap based on the correlation coefficients matrix in all three measures (i.e., average geodesic proximity, average temporal proximity, and average spatio-temporal proximity). The grids are colored based on obtained correlation coefficients, such that blue shows negative correlations, red shows positive correlations, and white shows no correlation among pairs of children, following the color bar.

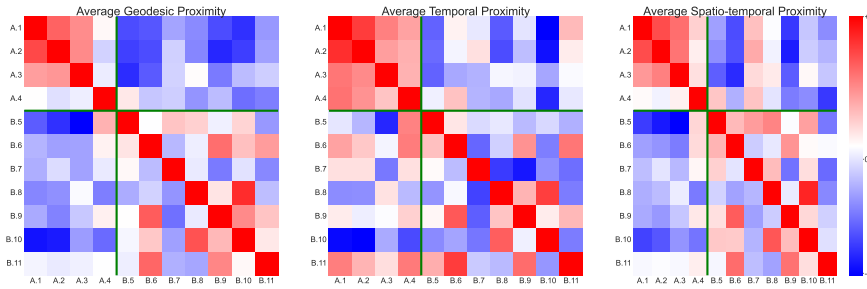
We further quantified and visualized the results of these heatmaps using the box plots to better understand the relations in each playgroup. Figure 3.5 shows the box plot of inter-group and cross-group correlation of the results per proximity measure in both playgroups.

In Figure 3.5-a, we observed that in PG-1-A, the highest correlation was found in pairwise temporal proximity, compared with the other two proximity measures. Meanwhile, in group PG-1-B, spatio-temporal proximity scored the highest correlation, indicating that spatial context could better describe the accessibility among children of this group. In addition, the average correlation is higher in PG-1-A compared with PG-1-B, among all proximity measures. This could be because of the smaller group size in PG-1-A. Overall, there is a positive correlation between the proximity measures of children from the same group and a negative correlation among children from different groups, except in the temporal proximity of PG-1-B, where the average cross-group correlation is higher than the inter-group correlation in the boxplot. This is also illustrated in Figure 3.4-a, where several negative correlations (cells in blue) were observed among children in PG-1-B. Nonetheless, no pronounced patterns between in-degree and out-degree proximity were found in this playgroup.

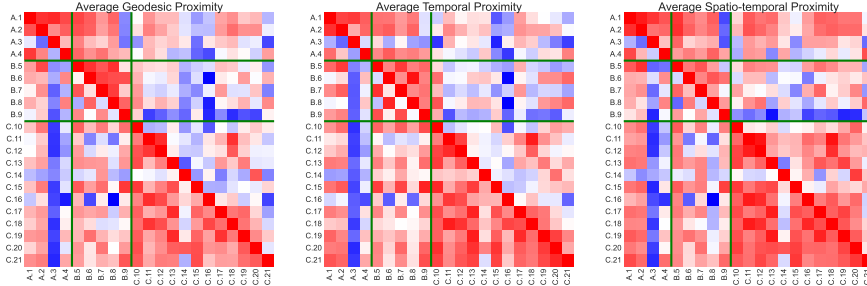
Figure 3.5-b shows that in PG-2 the positive and negative correlations are more uniformly distributed across inter-groups and cross-groups pairs. Yet, in all three groups (A, B, and C) and across all proximity measures, the average inter-group correlation is higher than the cross-group measures. This means that, in general, children are more accessible by peers from the same group, and it is easier to access peers from the same group in all three dimensions of geodesic, temporal, and spatio-temporal proximity. Yet, some children appear as outliers, not following the general pattern in the heatmap visualizations (e.g., A4 and B11 in PG-1, A3 and C14 in PG-2 as illustrated in Figure 3.4).

Furthermore, the triangular patterns are more pronounced in PG-2 than in PG-1. This shows a considerable difference between proximity towards a child from the peer group and proximity towards peers by a child. Specifically, in all three measures, children in PG-2-A and PG-2-B have a positive correlation in their in-degree

components (i.e., more red cells in their upper triangular heatmap). In comparison, in PG-2-C the positive correlation is observed in out-degree measures (more red cells in their lower triangular heatmap). This means that the peer networks in groups A and B had quicker access to a child, compared with the child accessing the network. On the contrary, in group C, individuals had quicker access to peer groups.



(a) PG-1 (groups A and B): the heatmap of pairwise accessibility in Geodesic proximity, temporal proximity and Spatio-temporal proximity.



(b) PG-2 (groups A, B, and C): the heatmap of pairwise accessibility in Geodesic proximity, temporal proximity and Spatio-temporal proximity.

Figure 3.4: The heatmap of Schoolyard Dataset for two playgroups: comparing the correlation coefficients matrix of geodesic proximity, temporal proximity, and spatio-temporal proximity among children in playgroup PG-1 and PG-2 (x.code = group id (A, B, or C). Child identifier code). The playgroup consists of several groups (A, B, and C) divided by green lines.

Furthermore, to better understand the impact of physical space in shaping identified clusters, we extracted the GPS location of children per cluster across one of the school breaks, and visualized the kernel density estimation (KDE) of the GPS locations. KDE represents the distribution of GPS locations using a continuous

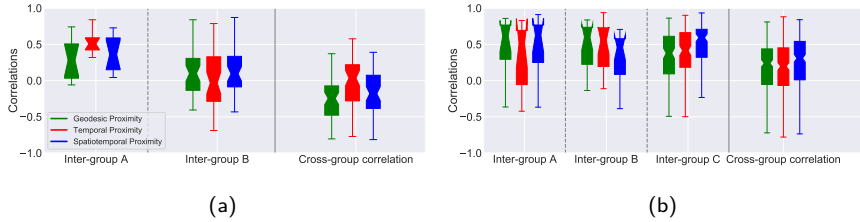


Figure 3.5: The correlation of pairwise proximity for two playgroups: comparing the CCM of geodesic proximity (in green), temporal proximity (in red), and spatio-temporal proximity (in blue) among children in playgroup PG-1 and PG-2. On average, children in both groups have a higher correlation with peers from the same group and a lower correlation with peers from different groups. This difference is more significant in PG-1-A and less significant in senior playgroup, i.e., PG-2.

probability density curve in two dimensions. The result of this analysis is depicted in Figure 3.6 and Figure 3.7 for PG-1 and PG-2, respectively.

As shown in Figure 3.6, we identified four clusters via hierarchical clustering of pairwise spatio-temporal proximity. Each cluster shows how the most accessible peers within each cluster utilized the playground. As a result, in *Cluster 1*, children were most accessible in the communal area where most play structures were quickly reachable (e.g., climbing frame, funnel ball, sandpit, and bench). Children in *Cluster 2* and *Cluster 4* often played in sandpits and quickly accessed their peers playing in the same location. According to field observations, children in PG-1 often attended to wheeled movements. This is observed in *Cluster 1* and *Cluster 3*, in which children widely used the playground via wheeled mobility toys (e.g., bikes, steppers, etc.). This finding shows how the available play structures impacted children's use of space, and could determine group activities in a network.

The impact of available play structures was even more evident in PG-2. In Figure 3.7, hierarchical clustering of pairwise spatio-temporal proximity identified seven clusters. In *Cluster 2*, with the most members, children often had face-to-face contact with peers around the play structures. This could, for instance, be due to using several play structures during the break (e.g., climbing frame, bench, and bar fixes) or being involved in an active game that required movements. Thus, the highly supplied area by the play structures attracted more children, and more conscious or unconscious contacts happened in this cluster. *Cluster 1* and *Cluster 3* belong to children who stayed around and in the climbing frame to climb, chat, or to simply observe other peers. The use of space as influenced by a specific play

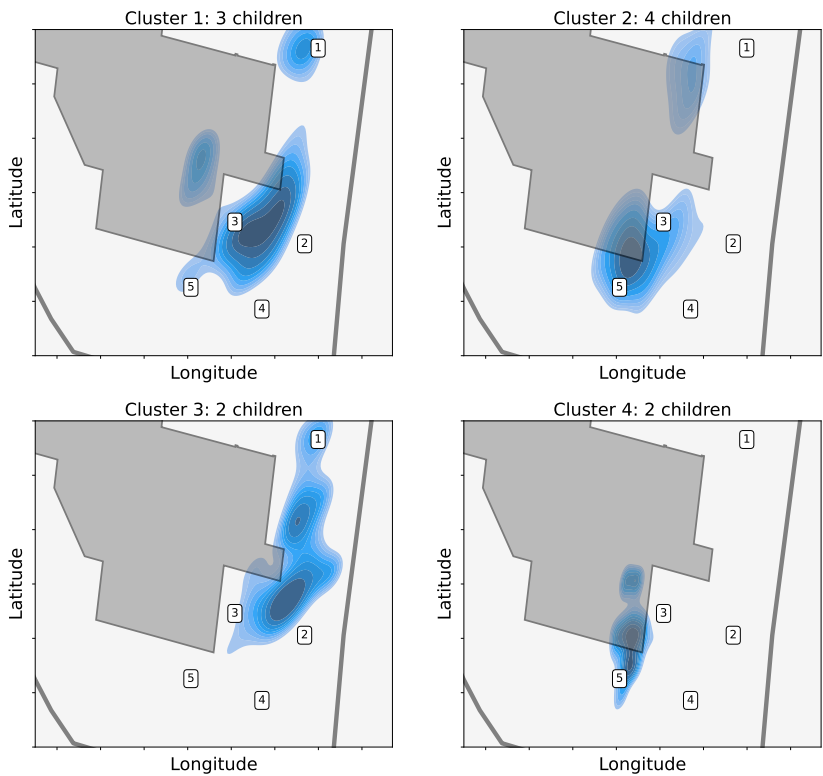


Figure 3.6: The GPS location of children in PG-1 is mapped to the school floor plan. The KDE plot in blue shows the spatial density of the GPS locations in the identified clusters, following the color bar.

3

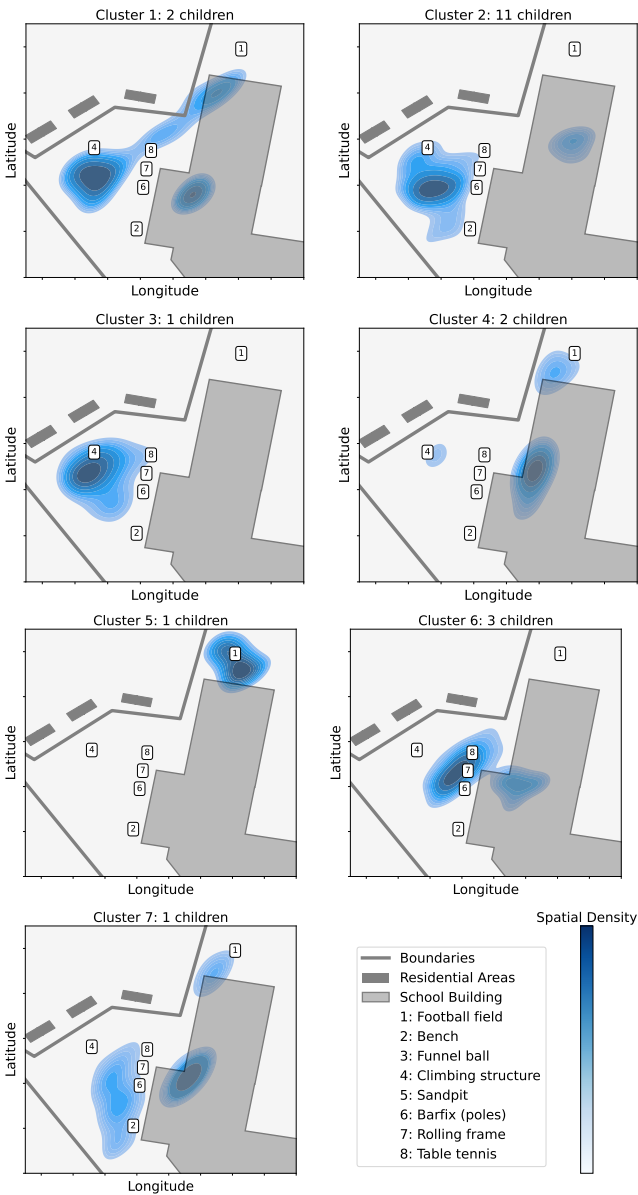


Figure 3.7: The GPS location of children in PG-2 is mapped to the school floor plan. The KDE plot in blue shows the spatial density of the GPS locations in the identified clusters, following the color bar.

structure was also observed in *Cluster 5* and *Cluster 6*, where the football field and the attached structures (i.e., table tennis, rolling game, and bar fix) were located. A higher number of clusters in this playgroup, compared with PG-1, occurred due to a higher number of participants in PG-2; more extensive physical space provided to this playgroup during recess; and differences in their age group, as senior grades tend to have more group-oriented activities, such as talking on a bench or table tennis, compared with junior grades.

3.6 Conclusions

The present study features a novel spatio-temporal proximity metric, which retains spatial information about the underlying temporal dynamics. The main strength of this metric, compared to the previously proposed temporal proximity metrics, lies in the ability to account for the role of the physical environment when analyzing an individual's temporal behavior. We applied the proposed metric to a case study, in an effort to understand children's social behavior in a schoolyard environment.

We evaluated the quality of the proposed metric by performing a downstream clustering task. Our results show that the proposed spatio-temporal metric uniquely quantified an individual's relationships over time, concerning the spatial context of their activity. This led to better clustering performance when compared to geodesic proximity metrics that were applied to senior groups in our study. In junior groups, the spatial context did not positively impact clustering performance.

We also introduced two new datasets collected from children playing in schoolyards during recess, using wearable proximity tags and GPS loggers. We made these datasets available for other researchers to study social dynamics. In this case study, we first examined pairwise accessibility based on the calculated proximity values among children in a playgroup. Our results showed that, in general, children were more accessible to peers from the same group. Similarly, children more easily accessed peers from the same group. Second, we analyzed the impact of the physical environment and its characteristics on forming the most accessible clusters. Our findings show that in both of the playgroups we studied, the most accessible clusters utilized common areas in the playground, which were often located around one or several play structures that the schoolyard featured.

In summary, we found that application of a novel spatio-temporal proximity metric to data collected via wearable proximity sensors offered valuable insight into the spatio-temporal accessibility of children in schoolyards. The proposed method could be used to develop and evaluate targeted interventions aimed at creating a more accessible and inclusive environment, for example, by designing new play

structures or interactive games aimed at increasing pairwise accessibility among all children. Moreover, the proposed method is applicable beyond schoolyards and can be used in diverse scenarios, including analyzing employee behavior in an office setting, tracking athletes' movements on sports fields, and monitoring the well-being of elderly individuals in a nursing home. The present study used the proposed metric in the context of a case study at an exploratory level. Yet, making broad generalizations about children's behavior requires a larger sample of data. Collecting a larger sample was not possible due to the COVID-19 crisis, which resulted in school closure and implementing strict protocols during school re-openings, but it currently remains one of the major focuses of the research team.

One of the limitations of this research is the absence of ground-truth data due to privacy concerns and the complexity of defining "truth" for the examined variables. Specifically, in order to provide ground-truth for children's spatio-temporal proximity, we need to know the details of their group behavior in time and space, e.g., the true values of children's pairwise meaningful interactions, the frequency of contacts and their accurate location in the schoolyard. Thus, it is important to consider some levels of uncertainty and the potential for alternative interpretations, e.g., uncertainties in sensors operations might conclude a different picture of schoolyard activities or use of some specific areas might be due to the impact of external events such as weather conditions, etc. Another limitation in the proposed method is the time complexity of shortest-path algorithms, which is $\mathcal{O}(V \log V + E)$ for the Dijkstra shortest-path algorithm, where E and V denote the number of edges and the number of nodes, respectively. Thus, it is computationally expensive to estimate the proximity metrics when the data is gathered from larger samples (higher number of users) and over a longer period of time (higher number of temporal edges). Employing shortest-path estimation algorithms and parallel computing is recommended to enhance the scalability of the proposed method for larger graphs in future research. Additional research is needed in order to understand factors contributing to an inclusive environment, to investigate how the proposed metrics describe such an environment, and to prioritize interventions using the proposed metrics in a recommender system. Moreover, future research might explore the Kinetic data structures, a data structure to track an attribute of a moving geometric system, to capture the spatio-temporal changes of moving users.